# Automatically Generated Icons for Documents of Electronic Libraries

Ádám Schmideg

Most documents on the internet assume users to have similar capabilities as those of computers. Keywords are listed, information is grouped in tree structures. Human mind however seems to work a different way. Everyone had an experience that he or she did not remember a sentence in a book but its location on the top of the page. Or he or she remembers a book's color and size but not its logical category in the library. The steady growth of information on the internet gave rise questions concerning document visualization. The emphasis however has been put on visualizing the structure of a collection of documents. The visual nature of human mind and memory suggests that separate documents also call for some kind of visualization or iconization. Since a text, even of moderate size, contains many ideas and thoughts, a detailed graphical representation would be very complex. Such a representation is likely to be less comprehensive than the original text. Thus, in order to get an emblematic icon for document

* one has to extract the content first
* find relatively simple icons for these concepts
* compose an image out of the basic icons

The first step is a recent and hot topic (though one approach was coined ICE - Intelligent Content Extraction) that is strongly related with the huge mass of information on the internet. The second step can be achieved using semantic knowledge to be able to select appropriate concepts as representatives of all possible thought within scope. And then it requires some artistic skill to depict these concepts.

We are mainly interested in the last step, i.e. in rules or algorithms of composing complex images from primitive ones. One approach is to use a template of placeholders for all images to be arranged. The simplest form of a template is a rectangular grid, so each image will be of the same size. More complicated patterns can be borrowed from heraldry where the constituting shapes are sometimes irregular, though the template as a whole has a delicate symmetry. Icons generated this way have very similar and easily predictable structure. In addition, they do not seem to capture some concepts in one but mostly fall apart into separate icons. Our method is to assign properties and special placeholders to each basic image. Then find a possible arrangement where all conditions are met - basic images fit into each other's placeholders, they are magnified to the extent allowed by their properties, etc.

This is a case of a complex packing problem that is difficult to solve in general. However, rules for placeholders and properties can be reduced to a subset. This new packing problem is much easier to solve and it still produces quite complex images closer to the ones generated manually. These icons are useful in many diverse areas,

* result set of a search on the Web
* visualization of heterogeneous databases
* collection of arbitrary hyperlinks
* intuitive user interfaces for internet applications

The generating method can be refined in many aspects, too.